Microsoft Azure - Machine Learning Process flow

For

Forecasting e-District Average Disposal Days

Prepared by:

MP State Agency for Public Services

1. Problem Statement

e-District portal is an online platform for delivering citizen services (such as Income certificate, Caste certificate, Revenue services etc.). It provides a channel to receive applications from citizens thus reduces time and cost in travelling to respective State Departments.

The applications received through e-District portal is forwarded to concerned officer of State Department who scrutinizes and processes the application and provide service to citizen in the form of Certificate, etc.

Under Public Service Guarantee Act, 2010, various citizen services have been notified which creates a charter for Officers to ensure delivery of services within timeline, so that citizens need not to re-visit State Departments time and again to ensure that their applications are processed, and services delivered to them within timeline.

While most of the services are delivered within timelines, some of them skip the timelines due to multiple reasons (delay in processing by Officer, incomplete application, incomplete documentation, wrong address etc.). **Hence, we are conducting an exercise to predict the average disposal days of any e-District application based on Service and Date of application.**
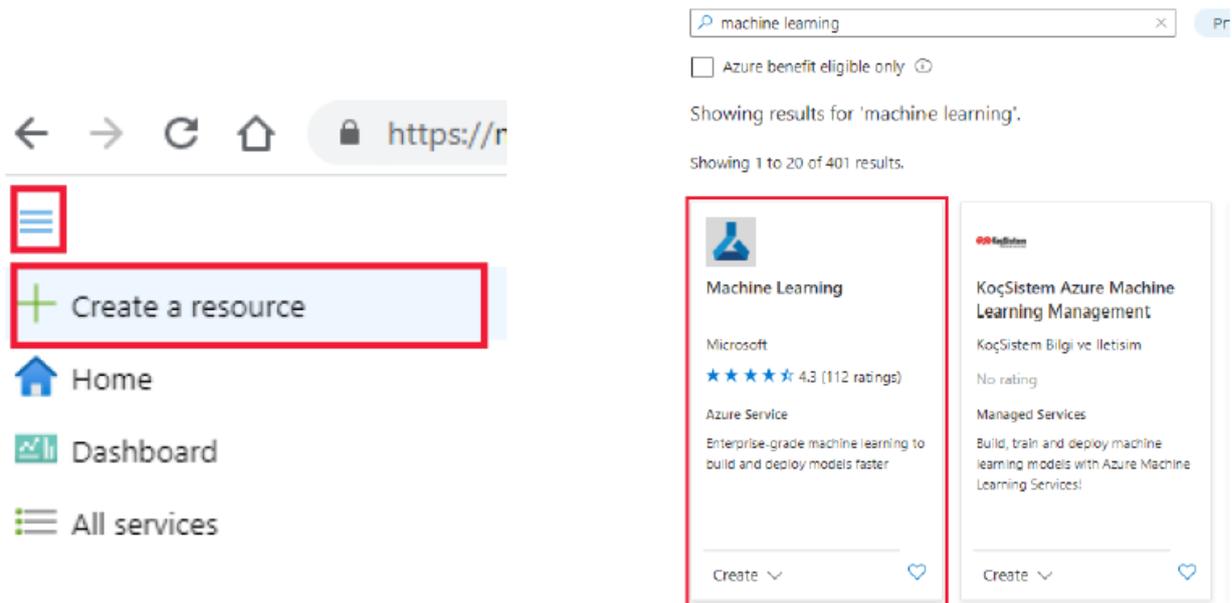
2. Data Source:

e-District database (Service name, Date, Average Disposal Days)

Service_Name_H	Date	Avg
(6.3-A) अनुसूचित जाति, अनुसूचित जनजाति के लिये जाति प्रमाण पत्र प्रदाय करना	2019-06-11T00:00:00Z	53
RCMS - 4.13 भूमि का सीमांकन	2018-11-09T00:00:00Z	82
मध्य प्रदेश वन मित्र दावा पंजीकरण(जनजातीय कार्य विभाग)		0
(12.3) दीनदयाल अंत्योदय उपचार योजना कार्ड जारी करना	2019-04-07T00:00:00Z	9
(2.4) निर्माण श्रमिकों का पंजीयन	2018-09-08T00:00:00Z	37
6.6(A) ऐसे नागरिक जिनके परिवार के किसी सदस्य (पिता/ भाई/बहन) को पूर्व में अनुविभागीय अधिकारी (राजस्व) द्वारा अनुसूचित जाति, अनुसूचित जनजाति रागाा ाग ाली किग		19

3. Process Workflow

3.1 CREATE A WORKSPACE

1. Sign into the Azure Portal using your Microsoft credentials.
2. Select +**Create** a resource, search for Machine Learning, and create a new Machine Learning resource.

3. After configuring the Workspace, select Review + Create.
4. Select Create, to create Workspace.
5. To view new Workspace, select Go to Resource.
6. From Portal view of Workspace, Select Launch Studio to go to the Azure ML Studio.
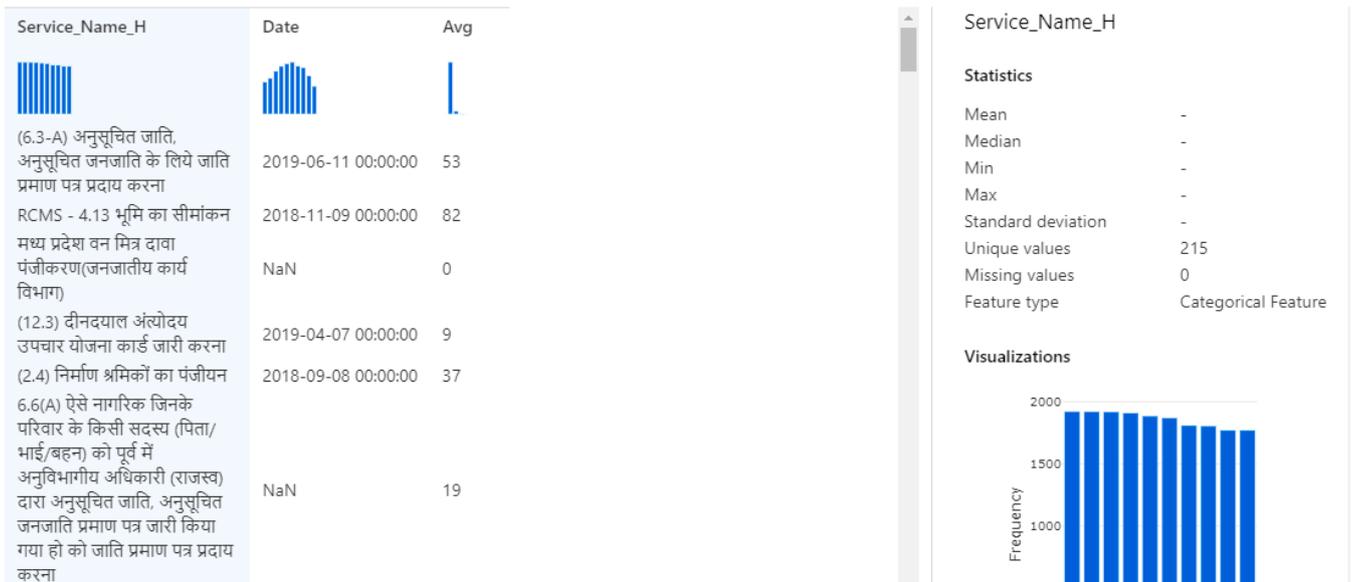
3.2 CREATE COMPUTE RESOURCES:

1. Create Compute Targets: On the Compute Instances tab, add a new compute instance:
2. Virtual Machine type: CPU
3. Virtual Machine size: Standard_DS11_v2 (Choose Select from all options to search for and select this machine size)
4. Compute name: cmhelpline1
5. SSH access: Unselected
6. While the compute instance is being created, switch to the Compute Clusters tab, and add a new compute cluster.

3.3 Import Data in Azure ML

Service_Name_H	Date	Avg
(6.3-A) अनुसूचित जाति, अनुसूचित जनजाति के लिये जाति प्रमाण पत्र प्रदाय करना	2019-06-11T00:00:00Z	53
RCMS - 4.13 भूमि का सीमांकन	2018-11-09T00:00:00Z	82
मध्य प्रदेश वन मित्र दावा पंजीकरण(जनजातीय कार्य विभाग)		0
(12.3) दीनदयाल अंत्योदय उपचार योजना कार्ड जारी करना	2019-04-07T00:00:00Z	9
(2.4) निर्माण श्रमिकों का पंजीयन	2018-09-08T00:00:00Z	37
6.6(A) ऐसे नागरिक जिनके परिवार के किसी सदस्य (पिता/भाई/बहन) को पूर्व में अनुविभागीय अधिकारी (राजस्व) दारा अनुसूचित जाति, अनुसूचित		19

3.4　　　Data Manipulation (if any)

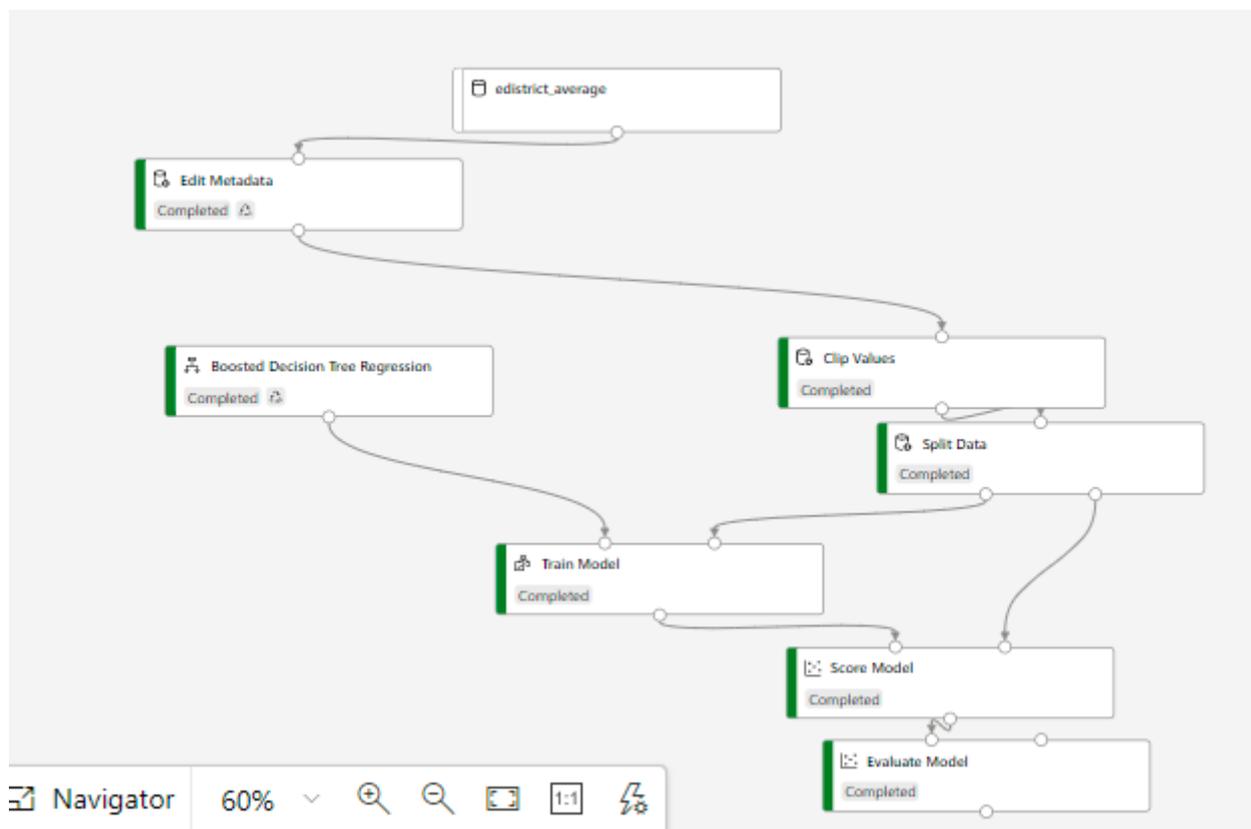　　　a)　Converted 'e-District services' from 'String' to 'Categorical' variable

Service_Name_H	Date	Avg
(6.3-A) अनुसूचित जाति, अनुसूचित जनजाति के लिये जाति प्रमाण पत्र प्रदाय करना	2019-06-11 00:00:00	53
RCMS - 4.13 भूमि का सीमांकन	2018-11-09 00:00:00	82
मध्य प्रदेश वन मित्र दावा पंजीकरण(जनजातीय कार्य विभाग)	NaN	0
(12.3) दीनदयाल अंत्योदय उपचार योजना कार्ड जारी करना	2019-04-07 00:00:00	9
(2.4) निर्माण श्रमिकों का पंजीयन	2018-09-08 00:00:00	37
6.6(A) ऐसे नागरिक जिनके परिवार के किसी सदस्य (पिता/भाई/बहन) को पूर्व में अनुविभागीय अधिकारी (राजस्व) दारा अनुसूचित जाति, अनुसूचित जनजाति प्रमाण पत्र जारी किया गया हो को जाति प्रमाण पत्र प्रदाय करना	NaN	19

Service_Name_H

Statistics

Mean	-
Median	-
Min	-
Max	-
Standard deviation	-
Unique values	215
Missing values	0
Feature type	Categorical Feature

Visualizations

　　　b)　Clip Values: To remove outliers

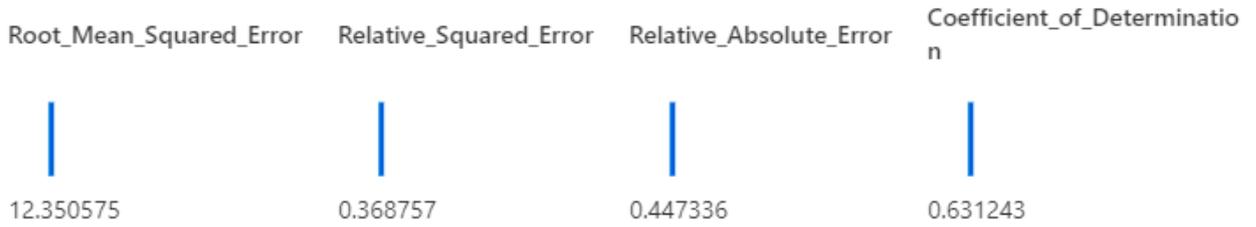| Service_Name_H	Date	Avg
(6.3-A) अनुसूचित जाति, अनुसूचित जनजाति के लिये जाति प्रमाण पत्र प्रदाय करना	2019-06-11 00:00:00	53
RCMS - 4.13 भूमि का सीमांकन	2018-11-09 00:00:00	82
मध्य प्रदेश वन मित्र दावा पंजीकरण(जनजातीय कार्य विभाग)	NaN	0
(12.3) दीनदयाल अंत्योदय उपचार योजना कार्ड जारी करना	2019-04-07 00:00:00	9
(2.4) निर्माण श्रमिकों का पंजीयन	2018-09-08 00:00:00	37
6.6(A) ऐसे नागरिक जिनके परिवार के किसी सदस्य (पिता/भाई/बहन) को पूर्व में अनुविभागीय अधिकारी (राजस्व) दारा अनुसूचित जाति, अनुसूचित जनजाति प्रमाण पत्र जारी किया गया हो को जाति प्रमाण पत्र प्रदाय करना	NaN	19

3.5 Machine Learning Algorithm used

We have used Boosted Decision Tree Regression method

4. Achievement/ Result

The coefficient of determination value is (0.631)

Root_Mean_Squared_Error	Relative_Squared_Error	Relative_Absolute_Error	Coefficient_of_Determination
12.350575	0.368757	0.447336	0.631243

5. Model Deployment

The model shall be deployed on e-District portal for citizens to predict average disposal days of e-District services.

Input data to test real-time endpoint   **Test**

Test result   parsed | raw

⌄ WebServiceInput0

Service_Name_H

(6.3-A) अनुसूचित जाति, अनुसूचित जनजाति के लिये जाति प्रमाण पत्र ...

Date

2019-06-11T00:00:00Z

Avg

53

⌄ WebServiceOutput0

key	value
Service_Name_H	(6.3-A) अनुसूचित जाति, अनुसूचित जनजाति के लिये जाति प्रमाण पत्र प्रदाय कर
Date	2019-06-11 00:00:00
Avg Disposal Days (Predicted)	47.3